WHAT IS CLAIMED IS:

1.   A method for matching voice characteristics of a disc jockey, said method comprising:

receiving, by a sound characteristic estimator, a first segment of audio signal;

5   determining, by said sound characteristic estimator, a first set of sound characteristics from said first segment of audio signal;

receiving, by said sound characteristic estimator, a second segment of audio signal;

determining, by said sound characteristic estimator, a second set of sound characteristics from said second segment of audio signal; and

10   interpolating a voice characteristic transition for said disc jockey from said first set of sound characteristics to said second set of sound characteristics between a starting time and an ending time.

2.   The method according to claim 1, wherein said first segment of audio signal includes an audio signal of a song.

15   3.   The method according to claim 1, wherein said first segment of audio signal includes an audio signal of a sports program.

4.   The method according to claim 1, wherein said sound characteristics include pitch.

5.   The method according to claim 1, wherein said sound characteristics include tempo.

20   6.   The method according to claim 1, wherein said sound characteristics include volume.

7.   The method according to claim 1, wherein said interpolating comprises:

converting said first set and said second set of sound characteristics of said segments of audio signals to a corresponding first set of voice characteristics and second set of voice characteristics of said disc jockey; and

generating an interpolation between said first set of voice characteristics and said

5   second set of voice characteristics of said disc jockey to produce said voice characteristics transition.

8.   The method according to claim 7, wherein said generating an interpolation includes generating said interpolation using a linear method.

9.   The method according to claim 7, wherein said generating an interpolation

10   includes generating a voice transition between a voice characteristic from said first set of voice characteristics and a corresponding voice characteristic from said second set of voice characteristics.

10.   The method according to claim 7, wherein said voice characteristics include average pitch.

15   11.   The method according to claim 7, wherein said voice characteristics include speaking rate.

12.   The method according to claim 7, wherein said voice characteristics include loudness.

13.   The method according to claim 7, wherein said voice characteristics include

20   prosody.

14.   The method according to claim 1, further comprising:

receiving, by a synthetic disc jockey, a piece of text, said voice characteristic transition, said starting time, and said ending time; and

generating, by said synthetic disc jockey using a text-to-speech engine, a speech signal with a duration from said starting time to said ending time based on said piece of text and said voice characteristic transition.

15. The method according to claim 14, further comprising choosing a sample set of voice characteristics for said synthetic disc jockey based on a genre of said first segment of audio signal.

16. The method according to claim 14, wherein said piece of text represents announcement information of a disc jockey.

17. The method according to claim 14, further comprising rendering said speech signal to generate an announcement of said synthetic disc jockey.

18. A computer-readable medium encoded with a plurality of processor-executable instruction sequences for:

receiving, by a sound characteristic estimator, a first segment of audio signal;

determining, by said sound characteristic estimator, a first set of sound characteristics from said first segment of audio signal;

receiving, by said sound characteristic estimator, a second segment of audio signal;

determining, by said sound characteristic estimator, a second set of sound characteristics from said second segment of audio signal; and

interpolating a voice characteristic transition for said disc jockey from said first set of sound characteristics to said second set of sound characteristics between a starting time and an ending time.

19. The computer-readable medium according to claim 18, wherein said first segment of audio signal includes an audio signal of a news program.

20. The computer-readable medium according to claim 18, wherein said sound characteristics include tempo.

21. The computer-readable medium according to claim 18, wherein said interpolating comprises:

converting said first set and said second set of sound characteristics of said segments of audio signals to a corresponding first set of voice characteristics and second set of voice characteristics of said disc jockey; and

generating an interpolation between said first set of voice characteristics and said second set of voice characteristics of said disc jockey to produce said voice characteristics transition.

22. The computer-readable medium according to claim 21, wherein said generating an interpolation includes generating said interpolation using a linear method.

23. The computer-readable medium according to claim 21, wherein said generating an interpolation includes generating a voice transition between a voice characteristic from said first set of voice characteristics and a voice characteristic from said second set of voice characteristics.

24. The computer-readable medium according to claim 21, wherein said voice characteristics include dynamic range of pitch.

25. The computer-readable medium according to claim 18, said computer-readable medium being further encoded with processor-executable instruction sequences for:

receiving, by a synthetic disc jockey, a piece of text, said voice characteristic transition, said starting time, and said ending time; and

generating, by said synthetic disc jockey using a text-to-speech engine, a speech signal with a duration from said starting time to said ending time based on said piece of text and said voice characteristic transition.

26.   A system for matching voice characteristics of a disc jockey, said system comprising:

a sound characteristic estimator, said estimator being configured to receive a first and a second segment of audio signal, and to respectively determine a first and a second set of sound characteristics from said first and second segments of audio signal; and

an interpolator, said interpolator being configured to interpolate a voice characteristic transition for said disc jockey from said first set of sound characteristics to said second set of sound characteristics between a starting time and an ending time.

27.   The system according to claim 26, wherein said sound characteristics include pitch.

28.   The system according to claim 26, wherein said interpolator is configured to:

convert said first set and said second set of sound characteristics of said segments of audio signals to a corresponding first set of voice characteristics and second set of voice characteristics of said disc jockey; and

generate an interpolation between said first set of voice characteristics and said second set of voice characteristics of said disc jockey to produce said voice characteristics transition.

29.   The system according to claim 28, wherein said interpolator generates an interpolation using a linear method.

30.   The system according to claim 26, further comprising:

a synthetic disc jockey, said synthetic disc jockey being configured to receive a piece of text and said voice characteristic transition; and

a text-to-speech engine,

wherein said synthetic disc jockey is configured to generate, using said text-to-speech engine, a speech signal with a duration from said starting time to said ending time based on said piece of text and said voice characteristic transition.